

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kanker payudara merupakan salah satu kanker paling umum yang dialami oleh wanita. Terhitung sebanyak 39.520 orang tercatat meninggal akibat kanker payudara diperkirakan terjadi pada wanita Amerika Serikat pada tahun 2011 [1]. *Invasive ductal carcinoma* merupakan salah satu jenis kanker payudara yang sering terjadi. Hampir 70-80% dari semua diagnosis kanker payudara. Menurut American Cancer Society, setiap tahunnya terdapat lebih dari 18.000 wanita di negara Amerika Serikat, diketahui bahwa mereka menderita kanker payudara jenis *invasive*. Sebagian besar dari mereka didiagnosis *invasive ductal carcinoma* [1]. Inilah alasan mengapa IDC termasuk masalah yang mengkhawatirkan bagi dunia jika tingkat kelangsungan hidupnya rendah.

Beberapa metode yang digunakan untuk mendeteksi payudara terdapat mammografi dan MRI. Tetapi, sejauh ini metode yang paling umum digunakan adalah skrining mamografi. Hal itu disebabkan karena banyak negara yang memiliki inisiatif skrining untuk kanker payudara [2]. Mamografi telah menjalani pemeriksaan yang lebih besar daripada hampir semua intervensi medis lainnya, terutama karena ketika skrining mamografi dimasukkan ke dalam populasi, kematian akibat kanker payudara menurun [3].

Dalam melakukan penelitian tentang kanker, *deep learning* dengan metode *convolutional neural network* paling banyak digunakan. Khususnya pada hal mendeteksi dan klasifikasi kanker payudara [4]. Kanker payudara jenis *invasive ductal carcinoma* telah dilakukan oleh beberapa peneliti dalam penerapan deteksi objek dengan *convolutional neural network*.

Pemrosesan citra otomatis untuk mendeteksi kanker payudara pada histopatologi telah menjadi topik penelitian selama beberapa waktu. Hal ini menjadi tantangan bagi para peneliti untuk mendapatkan akurasi yang baik untuk deteksi *invasive ductal carcinoma*. Deteksi kanker payudara dengan jenis *invasive ductal carcinoma* telah dieksplorasi oleh peneliti lain sebelumnya. Penelitian yang dilakukan oleh Romano,dkk dengan penelitian peningkatan pendekatan *deep learning* untuk memprediksi *invasive ductal carcinoma* dari gambar histopatologi.

Menggunakan ekstraksi fitur CNN dengan pooling layer baru yaitu *accept-reject pooling*. Dataset yang digunakan sebanyak 50.846, dengan splitting data sebanyak 40.672 untuk data train dan 10.168 untuk data test. Hasil eksperimen mampu meraih f-score 85,28% dan akurasi seimbang 85,41% dengan peningkatan dari 11,51% pada f-score dan 0,86% pada akurasi seimbang [5]. Penelitian lain oleh Rahman,dkk mengusulkan metode baru menggunakan *deep learning* yaitu CNN untuk klasifikasi IDC. Dataset yang digunakan pada penelitian tersebut sebanyak 157.572 dengan jumlah 126.064 sebagai data train dan 31.508 sebagai data test. Hasil eksperimen didapatkan tingkat akurasi 89% dan f1-score 89,34% [6]. Selain itu penelitian yang dilakukan oleh Romeno,dkk dalam klasifikasi sel *Invasif Ductal Carcinoma* (IDC) dalam slide histopatologi dengan model yang dibuat dari arsitektur Inepsi, mengusulkan modul normalisasi batch multi-level antara setiap langkah konvolusional dengan metode CNN. Jumlah dataset yang digunakan sebanyak 227.522, dengan splitting data sebanyak 94.543 untuk data train, 31.514 untuk data validasi, dan 151.465 untuk datatest. Hasil eksperimen didapatkan tingkat akurasi seimbang dan f-score masing-masing 0,890 dan 0,897 [7].

Pada penelitian ini, metode yang diusulkan menggunakan *deep learning* dengan ekstraksi fitur CNN. Beberapa peneliti mengungkapkan bahwa deep learning telah menunjukkan kemampuan yang sangat baik dalam mempelajari fitur / ciri-ciri pencitraan [8][9]. Metode yang digunakan untuk mengklasifikasi *invasive ductal carcinoma* positif dan *invasive ductal carcinoma* negatif dari citra histopatologi menggunakan model sudah ada yaitu VGG16 dan MobileNet. Model arsitektur VGG16 memiliki potensi yang bagus dalam diagnosis diferensial karsinoma tiroid papiler dalam gambar sitologis [10]. Selain VGG16, MobileNet juga memiliki performa yang cukup baik dalam klasifikasi citra [11]. Kontribusi pada penelitian ini adalah optimalisasi hyperparameter pada model MobileNet dan VGG16 menggunakan tuning hyperparameter yang terlibat dalam kinerja metode guna mengurangi loss dan mencegah overfitting. Selain itu, penerapan fine-tuning yang merupakan konsep transfer learning pada kedua model [12]. Hasil akhir akan dibandingkan dengan penelitian sebelumnya dan metode yang telah diusulkan diharapkan mencapai tingkat akurasi yang lebih baik dari hasil penelitian sebelumnya [5].

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah, adapun rumusan masalah pada penelitian ini antara lain:

- a. Bagaimana menentukan model yang lebih baik dari penelitian yang dilakukan sebelumnya untuk klasifikasi *Invasive Ductal Carcinoma* (IDC) pada citra histopatologi menggunakan *Convolutional Neural Network* (CNN) ?
- b. Bagaimana perbandingan akurasi antara MobileNet dan VGG16 dengan penelitian sebelumnya ?

1.3 Tujuan Penelitian

Tujuan dari penelitian ini mendapatkan hasil akurasi yang lebih baik dari penelitian yang dilakukan sebelumnya dari model yang diusulkan untuk klasifikasi *Invasive Ductal Carcinoma* (IDC) pada citra histopatologi menggunakan *Convolutional Neural Network* (CNN), serta membandingkan hasil dari model MobileNet dan Vgg16 dengan penelitian sebelumnya.

1.4 Batasan Masalah

Terdapat beberapa batasan masalah agar pembahasan yang dilakukan bisa terfokus, antara lain :

- a. Pembangunan system klasifikasi menggunakan bahasa pemrograman Phyton
- b. Dataset yang digunakan adalah data citra histopatologi yang diambil dari situs Kaggle (<https://kaggle.com>) yang berisi sebanyak 198.738 IDC (-) dan 78.786 IDC (+).
- c. Jumlah dataset yang digunakan pada proses penelitian sebanyak 50.840 yang diambil secara random dan mengandung IDC negative dan IDC positif dengan jumlah yang seimbang.
- d. Pembagian data train dan data test masing-masing 80:20 dari total dataset.
- e. Hasil klasifikasi menjadi dua kategori yaitu IDC positif dan IDC negative.